

# ON A DIFFERENTIAL GEOMETRIC VIEWPOINT OF JAYNES' MAXENT METHOD AND ITS QUANTUM EXTENSION

S. A. Ali<sup>1,2</sup>, Carlo Cafaro<sup>3</sup>, Adom Giffin<sup>4</sup>, Cosmo Lupo<sup>3</sup>, Stefano Mancini<sup>3</sup>

<sup>1</sup>*International Institute for Theoretical Physics and Mathematics Einstein-Galilei, via Santa Gonda 14, 59100 Prato, Italy*

<sup>2</sup>*Department of Arts and Sciences, Albany College of Pharmacy and Health Sciences,  
106 New Scotland Avenue, Albany, NY 12208, USA*

<sup>3</sup>*School of Science and Technology, Physics Division,  
University of Camerino, I-62032 Camerino, Italy*

<sup>4</sup>*Princeton Institute for the Science and Technology of Materials,  
Princeton University, Princeton, NJ 08540, USA*

We present a differential geometric viewpoint of the quantum MaxEnt estimate of a density operator when only incomplete knowledge encoded in the expectation values of a set of quantum observables is available. Finally, the additional possibility of considering some prior bias towards a certain density operator (the prior) is taken into account and the unsolved issues with its quantum relative entropic inference criterion are pointed out.

PACS numbers: Probability Theory (02.50.Cw), Riemannian Geometry (02.40.Ky), Entropy (89.70.Cf).

## I. INTRODUCTION

It is well established that *geometry* plays an important role in characterizing and understanding both classical and quantum physics. After all, it has been an old dream to reduce the fundamental laws of physics to geometry since Einstein's formulation of general relativity. In particular, it is a remarkable achievement that all the building blocks of quantum field theory can be formulated in terms of geometric concepts such as vector bundles, connections, curvatures, covariant derivatives and spinors [1].

It is well known that all practical problems in science are characterized by incompletely specified situations where we can only provide an *inference* (plausible conjecture) [2]: given incomplete information about reality, we choose probability distributions that represent the state of (incomplete) knowledge of the system being considered. The incompleteness of information leads to the non-uniqueness of the reconstructibility of the state of the classical system, that is, more than one situation is compatible with what is known. However, some situations are more likely to occur than others. This leads to a probabilistic description of dynamical systems in the presence of partial knowledge. Furthermore, it is also the case that geometry can be of some use for the study of inference. In particular, a number of investigations have shown that a useful approach to the study of statistical inference is to regard a parametric statistical model as a differentiable manifold equipped with a metric structure [3]. Of course, as pointed out by Skilling [4], like any other professional tool, "*geometry should be used with intelligence and care*".

In recent years, we have employed information geometry [5] (Riemannian geometry applied to probability theory) and inductive inference (the Maximum relative Entropy method (MrE) [6] of which Jaynes' MaxEnt [7] is a special case) to study the complexity of informational geodesic flows on curved statistical manifolds (statistical models) underlying the probabilistic description of physical systems in the presence of incomplete information [8]. Most of this work has dealt with ordinary probability distributions. However, what happens if the state of the system is represented by a quantum density operator? Are Riemannian geometric techniques of some utility in such cases? Is there an inference method useful to assign a density operator given a specific set of expectation values of quantum observables? Is there a quantum entropic inference method that can be used to update density operators when some prior bias towards a certain density operator (the prior) is assumed? Such inference problems occur in situations in which information is gained in two steps. In the first step, experiments are carried out to determine a prior density operator  $\rho_0$ . In the second step, more knowledge is gained about the expectation values of certain quantum observables  $\{A_k\}$ . How should one determine the density operator  $\rho$  which is best suited to make further predictions, accounting fully for both types of information? In this work we provide a few answers to such questions and emphasize some conceptual and computational problems that may emerge.

## II. INFERENCE AND DENSITY OPERATORS

Jaynes' MaxEnt method can be extended to the quantum density-matrix formalism in a straightforward manner [9].

Assume we are given the expectation values of the quantum operators  $A_1, \dots, A_m$ . Stated otherwise, the quantities  $A_j$  with  $j = 1, \dots, m$  are the quantum observables which represent the slow variables of the theory [10], that is, those whose expectation values can be measured at any given time. For the sake of reasoning, assume that the information constraints are given by,

$$\bar{A}_j = \langle A_j \rangle \stackrel{\text{def}}{=} \text{tr}(\rho A_j), \text{ with } j = 1, \dots, m. \quad (1)$$

We point out that if the constraints (1) involve only commuting observables  $A_j$ , the quantum method can be reduced to the classical one. In this case, there exists a basis in the Hilbert space in which all  $A_j$  are represented by diagonal matrices. In this basis,  $\rho$  should also be diagonal and diagonal elements are ordinary probabilities. Indeed, it can be shown that Jaynes' MaxEnt method also works for non-commuting observables [11]. In this case, however, the off-diagonal elements of  $\rho$  must be considered and this has no equivalent in standard probability theory. The values  $\bar{A}_j$  are averages obtained in a large number of experiments performed on an ensemble of identically prepared systems. The necessity of an ensemble is obvious if we consider a situation in which the state to be characterized is specified by giving the expectation values of non-commuting observables. This requires that measurements associated with each (non-commuting) observable is performed on different samples, belonging to separate sub-ensembles of the full ensemble whose state is described by  $\rho$ . Using the Lagrange multipliers technique, it can be shown that the density matrix  $\rho$  which represents the most unbiased picture of the state of the system on the basis of this much information is the one maximizing the von Neumann entropy  $\mathcal{S}(\rho)$ ,

$$\mathcal{S}(\rho) \stackrel{\text{def}}{=} -\text{tr}(\rho \log \rho), \quad (2)$$

subject to the information constraints in (1) and to the normalization condition  $\text{tr}(\rho) = 1$ . As a side remark, we point out that  $\mathcal{S}(\rho)$  is invariant under unitary transformations. This is an important property of the von Neumann entropy and is very desirable since we require that our macroscopic predictions should remain unchanged if the observables are unitarily transformed. As in the classical case, the maximization process is accomplished by finding the density matrix  $\hat{\rho}$  which solves the variational problem,

$$\delta \left[ -\text{tr}(\rho \log \rho) - \lambda_0 (\text{tr} \rho - 1) - \sum_{k=1}^m \lambda_k (\text{tr}(\rho A_k) - \bar{A}_k) \right] = 0, \quad (3)$$

where  $\delta = \delta_\rho$  is the variational operator. After some simple algebra, it follows that Eq. (3) reads,

$$(\delta \rho) \left[ \log \rho + \lambda_0 \mathbf{1} + \sum_{k=1}^m \lambda_k A_k \right] + \rho \left( \delta \left[ \log \rho + \lambda_0 \mathbf{1} + \sum_{k=1}^m \lambda_k A_k \right] \right) = 0, \quad (4)$$

where  $\mathbf{1}$  is the identity operator. We point out that one of the main technical difficulties that occur in the quantum setting is that the chain rule for derivatives does not hold:  $\rho$  and  $\delta \rho$  do not necessarily commute [10]. That said, the solution of the variational problem in (4) is achieved by imposing the condition,

$$\log \rho + \lambda_0 \mathbf{1} + \sum_{k=1}^m \lambda_k A_k = 0, \quad (5)$$

i.e.,

$$\rho = \exp \left[ -\lambda_0 \mathbf{1} - \sum_{k=1}^m \lambda_k A_k \right]. \quad (6)$$

The Lagrange multipliers  $\lambda_0$  in (6) is determined by the normalization condition  $\text{tr}(\hat{\rho}) = 1$ . As a matter of fact, noticing that  $\exp(-\lambda_0 \mathbf{1})$  equals  $\exp(-\lambda_0) \mathbf{1}$ , the condition  $\text{tr}(\rho) = 1$  yields

$$1 = \text{tr} \left[ \exp \left( -\lambda_0 \mathbf{1} - \sum_{k=1}^m \lambda_k A_k \right) \right] = \exp(-\lambda_0) \text{tr} \left[ \exp \left( -\sum_{k=1}^m \lambda_k A_k \right) \right], \quad (7)$$

that is,

$$\lambda_0 = \log \mathcal{Z}(\lambda_1, \dots, \lambda_m), \quad (8)$$

where  $\mathcal{Z}$  is the so-called partition function defined as,

$$\mathcal{Z} = \mathcal{Z}(\lambda_1, \dots, \lambda_m) \stackrel{\text{def}}{=} \text{tr} \left[ \exp \left( - \sum_{k=1}^m \lambda_k A_k \right) \right]. \quad (9)$$

The remaining Lagrange multipliers  $\lambda_k$  with  $k = 1, \dots, m$  are determined by the information constraints (1). In particular one finds,

$$\bar{A}_j = \langle A_j \rangle \stackrel{\text{def}}{=} \text{tr}(\rho A_j) = \exp(-\lambda_0) \text{tr} \left[ A_j \exp \left( - \sum_{k=1}^m \lambda_k A_k \right) \right] = \frac{1}{\mathcal{Z}(\lambda_1, \dots, \lambda_m)} \left( - \frac{\partial \mathcal{Z}(\lambda_1, \dots, \lambda_m)}{\partial \lambda_j} \right), \quad (10)$$

that is,

$$\bar{A}_j = - \frac{\partial \log \mathcal{Z}(\lambda_1, \dots, \lambda_m)}{\partial \lambda_j}. \quad (11)$$

Thus, the Lagrange multipliers  $\lambda_j$  with  $j = 1, \dots, m$  are implicitly defined in Eq. (11). Indeed, the multipliers  $\lambda_j$  may also be expressed in an explicit manner. Observing that  $S_{\max} = S(\rho)$  with  $\rho$  in (6) is given by,

$$\mathcal{S}_{\max} = \mathcal{S}_{\max}(\bar{A}_1, \dots, \bar{A}_m) = \lambda_0 + \lambda_1 \bar{A}_1 + \dots + \lambda_m \bar{A}_m, \quad (12)$$

we conclude that,

$$\lambda_j = \frac{\partial \mathcal{S}_{\max}(\bar{A}_1, \dots, \bar{A}_m)}{\partial \bar{A}_j}. \quad (13)$$

In general, it may be highly non trivial to solve explicitly the system of equations for the Lagrange multipliers  $\lambda_j$  in (13). Finally, substituting (8) and (13) into (6), we obtain the "quantum MaxEnt estimate",

$$\rho = \rho(A_1, \dots, A_m | \bar{A}_1, \dots, \bar{A}_m) = \frac{\exp \left[ - \sum_{k=1}^m \frac{\partial \mathcal{S}_{\max}(\bar{A}_1, \dots, \bar{A}_m)}{\partial \bar{A}_k} A_k \right]}{\text{tr} \left[ \exp \left( - \sum_{k=1}^m \frac{\partial \mathcal{S}_{\max}(\bar{A}_1, \dots, \bar{A}_m)}{\partial \bar{A}_k} A_k \right) \right]}. \quad (14)$$

We point out that the quantum MaxEnt estimate is always a physical state thanks to the canonical form of the density operator in (14). We recall that although Jaynes' MaxEnt was originally developed for assigning probability distributions [7], it can also be regarded as a special case of the MrE method when updating probability distributions from a uniform prior using the Gibbs-Shannon entropy [6]. In the quantum framework, Jaynes' MaxEnt is employed to assign density operators via maximization of the von Neumann entropy [9]. However, although at the moment there is no universal quantum MrE method for updating density operators, the quantum MaxEnt formalism may be viewed as a limiting case of a possible updating criterion from uniform prior density operators where inference is carried out by means of the logarithmic quantum relative entropy,

$$\mathcal{S}(\rho || \rho_0) \stackrel{\text{def}}{=} -\text{tr}[\rho(\log \rho - \log \rho_0)]. \quad (15)$$

When the prior  $\rho_0$  is uniform and proportional to the identity operator 1 acting on an  $n$ -dimensional Hilbert space,  $\mathcal{S}(\rho || \rho_0)$  reads

$$\mathcal{S}(\rho || \rho_0) = \mathcal{S}(\rho) - \log n. \quad (16)$$

Thus, maximizing (2) is equivalent to maximizing (16). We remark that the use of (15) as a suitable entropic tool for quantum inferences (updating density matrices from arbitrary priors) appears in [12] as well. For forthcoming utility, we assume a bias towards a uniform prior density operator  $\rho_0$ . Then, the maximization of the quantum logarithmic relative entropy (16) subject to a fixed information constraint  $\bar{A} = \langle A \rangle$  leads to the MaxEnt estimate,

$$\rho(\lambda) = \rho_0 \frac{e^{-\lambda A}}{\text{tr}(e^{-\lambda A} \rho_0)} = \frac{e^{-\frac{\lambda}{2} A} \rho_0 e^{-\frac{\lambda}{2} A}}{\text{tr}(e^{-\lambda A} \rho_0)}. \quad (17)$$

The Jaynes MaxEnt method has been successfully applied to partial (incomplete) reconstruction of density operators of quantum systems from the available measured mean values of the system's observables (single spin- $\frac{1}{2}$ , two correlated

spins- $\frac{1}{2}$ , Greenberger-Horne-Zeilinger states, vibrational states of neutral atoms, etc.) [13]. Unfortunately, Jaynes' MaxEnt quantum formalism is not without criticisms in the scientific community. Indeed, it has been claimed that quantum state estimates derived via the principle of MaxEnt are fundamentally different from those obtained via the quantum Bayes rule [14]. This seems understandable and it is somehow expected. By definition, "a probability is an abstract notion that represents the degree of plausibility of a proposition, subject to information regarding that proposition" [15]. Thus if one has different information one should come to different probabilities. However, just as Maximum relative Entropy successfully showed that one can use a universal method (MrE) for any type of classical information, it may be possible that a quantum version of MrE will do the same for the quantum case.

### III. GEOMETRY AND DENSITY OPERATORS

In 1985, Campbell showed that geometry can be introduced into probability calculus as follows [16]: for a fixed probability distribution, define the inner product of two random variables to be the expectation of the product of these variables. Differential geometry emerges when we consider varying the probability distribution, either directly or through changing parameters on which the distribution depends. Within such a geometric framework, the sets of probability distributions are viewed as differentiable manifolds, the random variables appear as vectors and the expectation values of random variables are replaced with inner products in tangent spaces to such manifolds of probabilities. In particular, the MaxEnt estimate of a probability distribution given a prior and the information constraints is found by following an integral curve through the prior which is orthogonal (in the Fisher metric on the simplex) to the hyperplane defined by the information constraint equation until the constraint is satisfied. In 1995, Braunstein and Caves extended Campbell's ideas to the quantum framework [17]. Here, following their analysis, we provide a differential geometric viewpoint of the (generalized) quantum MaxEnt estimate (17).

Consider the quantum analogue  $\mathcal{M}_{\vec{\rho}}$  of the probability simplex, the space of density operators  $\vec{\rho}$  written as vectors in  $\mathcal{L}(\mathcal{H})$ , the linear space of all linear operators on a  $n$ -dimensional Hilbert space  $\mathcal{H}$ ,

$$\mathcal{M}_{\vec{\rho}} \stackrel{\text{def}}{=} \left\{ \vec{\rho} \in \mathcal{L}(\mathcal{H}) : \vec{\rho} \stackrel{\text{def}}{=} \sum_{i,j=1}^n \rho^{ij} \vec{e}_{ij}, \vec{\rho} = \vec{\rho}^\dagger, \text{tr}(\vec{\rho}) = 1, \vec{\rho} \geq 0 \right\}. \quad (18)$$

The space  $\mathcal{M}_{\vec{\rho}}$  is an  $(n^2 - 1)$ -dimensional *real* manifold with complicated boundary. An arbitrary linear operator vector  $\vec{V}$  on  $\mathcal{H}$  can be decomposed in terms of an operator vector basis  $\vec{e}_{ij} \stackrel{\text{def}}{=} |i\rangle \langle j|$  with  $i, j = 1, \dots, n$  as follows,

$$\vec{V} = \sum_{i,j=1}^n \langle i | \vec{V} | j \rangle \vec{e}_{ij} = \sum_{i,j=1}^n V^{ij} \vec{e}_{ij}. \quad (19)$$

The tangent space at  $\vec{\rho}$  is an  $(n^2 - 1)$ -dimensional *real* vector space of traceless Hermitian operators  $\vec{T}$ ,

$$\vec{T} = \sum_{i,j=1}^n T^{ij} \vec{e}_{ij}, \text{tr}(\vec{T}) = 0. \quad (20)$$

The action of 1-forms  $\tilde{F}$  expanded in terms of the dual basis  $\tilde{\omega}^{ji} \stackrel{\text{def}}{=} |i\rangle \langle j|$ ,

$$\tilde{F} \stackrel{\text{def}}{=} \sum_{i,j=1}^n F_{ij} \tilde{\omega}^{ji}, \quad (21)$$

on density operators  $\vec{\rho}$  is defined as follows,

$$\tilde{F}(\vec{\rho}) \equiv \langle \tilde{F}, \vec{\rho} \rangle = \sum_{i,j,l,k=1}^n F_{ij} \rho^{lk} \langle \tilde{\omega}^{ji}, \vec{e}_{lk} \rangle = \sum_{i,j,l,k=1}^n F_{ij} \rho^{lk} \delta_l^j \delta_k^i = \sum_{i,j=1}^n F_{ij} \rho^{ji} = \text{tr}(\tilde{F} \vec{\rho}) \equiv \langle \tilde{F} \rangle. \quad (22)$$

Therefore, an Hermitian 1-form  $\tilde{F} = \tilde{F}^\dagger$  is an ordinary quantum observable with  $\langle \tilde{F}, \vec{\rho} \rangle = \langle \tilde{F} \rangle$ . A metric structure  $g_{\vec{\rho}}(\cdot, \cdot)$  on the manifold  $\mathcal{M}_{\vec{\rho}}$  can be introduced by defining the metric's action on a pair of 1-forms  $\tilde{A}$  and  $\tilde{B}$  as follows,

$$g_{\vec{\rho}}(\tilde{A}, \tilde{B}) \stackrel{\text{def}}{=} \left\langle \frac{\tilde{A}\tilde{B} + \tilde{B}\tilde{A}}{2} \right\rangle = \text{tr} \left[ \left( \frac{\tilde{A}\tilde{B} + \tilde{B}\tilde{A}}{2} \right) \vec{\rho} \right] = \text{tr} \left[ \frac{\tilde{A}}{2} (\vec{\rho}\tilde{B} + \tilde{B}\vec{\rho}) \right] = \langle \tilde{A}, \mathcal{R}_{\vec{\rho}}(\tilde{B}) \rangle, \quad (23)$$

where  $\mathcal{R}_{\vec{\rho}}(\vec{B})$  is the raising operator mapping 1-forms (lower covariant components) to vectors (upper contravariant components),

$$\mathcal{R}_{\vec{\rho}}(\vec{B}) \stackrel{\text{def}}{=} \frac{\vec{\rho}\vec{B} + \vec{B}\vec{\rho}}{2}. \quad (24)$$

Such a metric is formulated in terms of statistical correlations of quantum observables. Using the lowering operator  $\mathcal{L}_{\vec{\rho}}(\vec{A})$  mapping vectors to 1-forms,

$$\mathcal{L}_{\vec{\rho}}(\vec{A}) = \mathcal{R}_{\vec{\rho}}^{-1}(\vec{A}), \quad (25)$$

we can also define the action of the metric tensor  $g_{\vec{\rho}}(\cdot, \cdot)$  on a pair of vectors  $\vec{A}$  and  $\vec{B}$ ,

$$g_{\vec{\rho}}(\vec{A}, \vec{B}) \stackrel{\text{def}}{=} \langle \mathcal{L}_{\vec{\rho}}(\vec{A}), \vec{B} \rangle = \text{tr} [\vec{B} \mathcal{L}_{\vec{\rho}}(\vec{A})]. \quad (26)$$

The quantum line element  $ds^2$ ,

$$ds^2 = g_{\vec{\rho}}(d\vec{\rho}, d\vec{\rho}), \quad (27)$$

with  $d\vec{\rho}$  given by,

$$d\vec{\rho} = \sum_{j=1}^n dp^j |j\rangle \langle j| + id\theta \sum_{m, l=1}^n (p^m - p^l) h_{lm} |l\rangle \langle m|, \quad (28)$$

and with  $e^{id\theta h}$  an infinitesimal unitary transformation on the orthonormal basis that diagonalizes  $\vec{\rho}$ , reads

$$ds^2 = \sum_{k=1}^n \frac{(dp^k)^2}{p^k} + 2d\theta^2 \sum_{j \neq k} \frac{(p^j - p^k)^2}{(p^j + p^k)} |h_{jk}|^2. \quad (29)$$

Notice that the quantum line element (29) is identical to the distinguishability metric for density operators obtained by Braunstein and Caves by optimizing over all generalized quantum measurements for distinguishing among neighboring quantum states [18].

The notion of quantum "transport" can also be introduced in this geometric framework. Consider a surface defined by the relation  $\langle \Delta \tilde{A} \rangle = 0$ , where  $\Delta \tilde{A} \stackrel{\text{def}}{=} \tilde{A} - \langle \tilde{A}, \vec{\rho} \rangle \tilde{1}$  is the zero mean observable and  $\tilde{1}$  is the unit 1-form operator. Any tangent vector  $\vec{t}$  that lies on the surface  $\langle \Delta \tilde{A} \rangle = 0$  satisfies  $\langle \Delta \tilde{A}, \vec{t} \rangle = 0$ . This implies that a vector field  $\mathcal{R}_{\vec{\rho}}(\Delta \tilde{A})$  associated with the observable  $\Delta \tilde{A}$  is orthogonal to the surface  $\langle \Delta \tilde{A} \rangle = 0$  since,

$$g_{\vec{\rho}}(\mathcal{R}_{\vec{\rho}}(\Delta \tilde{A}), \vec{t}) = \langle \Delta \tilde{A}, \vec{t} \rangle = 0. \quad (30)$$

The most efficient way to construct a path that goes from the state  $\vec{\rho}_0$  to another state that has a different expectation value for  $\tilde{A}$  is by consistently moving in the direction orthogonal to the surfaces  $\langle \Delta \tilde{A} \rangle = 0$ . This navigation is accomplished by selecting a parametrized path  $\vec{\rho} = \vec{\rho}(\lambda)$  with  $\vec{\rho}(\lambda = 0) = \vec{\rho}_0$  along the vector field  $\mathcal{R}_{\vec{\rho}}(\Delta \tilde{A})$  such that,

$$\frac{d\vec{\rho}(\lambda)}{d\lambda} + \mathcal{R}_{\vec{\rho}}(\tilde{A} - \langle \tilde{A} \rangle \tilde{1}) = 0, \quad (31)$$

that is,

$$\vec{\rho}(\lambda) = \frac{e^{-\frac{\lambda}{2}\tilde{A}}\vec{\rho}_0 e^{-\frac{\lambda}{2}\tilde{A}}}{\text{tr}(e^{-\lambda\tilde{A}}\vec{\rho}_0)}. \quad (32)$$

The path (32) is reminiscent of a quantum exponential model of *symmetric* type [19]. In conclusion, from a differential geometric point of view, the (generalized) MaxEnt estimate (17) can be regarded as a quantum trajectory (32) that passes orthogonally through the family of surfaces of constant expectation value corresponding to some observable (1-form).

#### IV. FINAL REMARKS

In this article, we presented a differential geometric viewpoint of the quantum MaxEnt estimate of a density operator when only incomplete knowledge encoded in the expectation values of a set of quantum observables is available. The additional possibility of considering some prior bias towards a certain density operator (the prior) was taken into account by using MrE and the unsolved issues with its relative entropic inference criterion were pointed out.

Although we limited our analysis to uniform priors and commutative observables (non-commutative extensions of (17) are under investigation), we must point out that care is needed in handling the non-commutativity of quantum observables. Furthermore, Jaynes' quantum MaxEnt method deals solely with *assigning* density operators. In the most optimistic scenario, we could use this method also in the presence of uniform priors. However, in actual scenarios, the prior is not uniform and a quantum entropic *updating* method is required. Unfortunately, there is a lack of a direct axiomatic justification (consistency requirements) for the use of the logarithmic quantum relative entropy maximization criterion for entropic inferences [11]. We also stress that although the Maximum relative Entropy method for updating probabilities includes both Jaynes' MaxEnt and Bayes' rule [6], a similar result in the quantum framework is still missing. Finally, Bayesian quantum inferences for determining quantum states depend on the prior as well as measurement data, even in the limit of an infinite number of measurements [20]. Furthermore, we are also aware that there has been some controversy over the naturalness of relative entropy as a tool for quantum statistical inference [21]. As a matter of fact, when the quantum-mechanical concept of relative entropy is discussed from an information-theoretic point of view, it can be shown that not all definitions found in the literature are equally suitable for the purpose of statistical inference by entropy maximization [22].

In view of all the above-mentioned considerations, we believe that an axiomatic formulation of a "universal" quantum entropic updating methodology embracing the quantum Bayes' rule [14] as a special case should to be developed. However, just as the Maximum relative Entropy method did this for the classical case, it is our hope that a quantum version (with non-uniform priors) can be constructed.

#### Acknowledgments

C. Cafaro thanks Ariel Caticha for very helpful discussions. The research of C. Cafaro, C. Lupo and S. Mancini has received funding from the European Commission's Seventh Framework Programme (FP7/2007–2013) under grant agreements no. 213681.

- 
- [1] T. Frankel, "*The Geometry of Physics*", Cambridge University Press (1997).
  - [2] E. T. Jaynes, "*Macroscopic Prediction*", in *Complex Systems and Operational Approaches in Neurobiology, Physics, and Computers*, ed. by H. Haken, Springer-Berlin, (1985).
  - [3] D. C. Brody and L. P. Hughston, *Phys. Rev. Lett.* **77**, 2851 (1996).
  - [4] J. Skilling, *AIP Conf. Proc.* **954**, 39 (2007).
  - [5] S. Amari and H. Nagaoka, "*Methods of Information Geometry*", Oxford University Press (2000).
  - [6] A. Caticha and A. Giffin, *AIP Conf. Proc.* **872**, 31 (2006); A. Giffin and A. Caticha, *AIP Conf. Proc.* **954**, 74 (2007).
  - [7] E. T. Jaynes, *Phys. Rev.* **106**, 620 (1957).
  - [8] C. Cafaro, *Chaos, Solitons & Fractals* **41**, 886 (2009).
  - [9] E. T. Jaynes, *Phys. Rev.* **108**, 171 (1957).
  - [10] M. R. Grasselli, "*Classical and Quantum Information Geometry*", PhD Thesis, King's College London (2001).
  - [11] R. Balian and N. L. Balazs, *Annals of Physics* **179**, 97 (1987).
  - [12] A. Caticha, "*Notes on Quantum Relative Entropy*", unpublished, University at Albany - SUNY (2004).
  - [13] V. Buzek, G. Drobny, R. Derka, G. Adam and H. Wiedemann, *Chaos, Solitons & Fractals* **10**, 981 (1999).
  - [14] R. Schack, T. A. Brun and C. M. Caves, *Phys. Rev.* **A64**, 014305 (2001).
  - [15] A. Giffin, "*Maximum Entropy: The Universal Method for Inference*", Ph. D. Thesis, SUNY at Albany, Albany (New York-USA), <http://arxiv.org/abs/0901.2987> (2008).
  - [16] L. L. Campbell, *Inform. Sci.* **35**, 199 (1985).
  - [17] S. L. Braunstein and C. M. Caves, *Annals of the New York Academy of Sciences* **755**, 786 (1995).
  - [18] S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994).
  - [19] O. E. Barndorff-Nielsen, R. D. Gill and P. E. Jupp, *J. Roy. Soc.* **B65**, 775 (2003).
  - [20] C. A. Fuchs and R. Schack, *AIP Conf. Proc.* **1101**, 255 (2009).
  - [21] M. J. Donald, *Commun. Math. Phys.* **105**, 13 (1968).
  - [22] R. W. Benoist, J.-P. Marchand and W. Wyss, *Lett. Math. Phys.* **3**, 169 (1979).